DOCUMENT RESUME

ED 107 578 SO 008 397

| | |
|---|---|
| AUTHOR | Angrist, Shirley S.; And Others |
| TITLE | Experimental Evaluation of Family Curriculum Materials for High School Students. |
| SPONS AGENCY | Buhl Foundation, Pittsburgh, Pa. |
| PUB DATE | 75 |
| NOTE | 39p.; A paper prepared for the Annual Meeting of the American Educational Research Association (Washington, D.C., April 1975) |
| EDRS PRICE | MF-$0.76 HC-$1.95 PLUS POSTAGE |
| DESCRIPTORS | Adolescence; American Studies; *Cross Cultural Studies; Cultural Differences; *Curriculum Development; *Curriculum Evaluation; Educational Research; *Family (Sociological Unit); *Family Life Education; Family Role; Formative Evaluation; Inquiry Training; Projects; Secondary Education; Social Studies Units; Sociology; Youth |
| IDENTIFIERS | *Family Life Curriculum Project; Israel; Japan |

ABSTRACT

This paper describes two new family life curriculum development projects at Carnegie-Mellon University and presents the results of an evaluation of two mini-courses using a modified Solomon four-group experimental design. Based on historical, sociological, anthropological, and psychological research, the first unit presents family life in Japan and the Kibbutz, while the second emphasizes American adolescence. In both courses students examine the changing norms, roles, role clusters, and socialization processes of family members in historical and contemporary contexts. The experimental design includes analyses of variance of posttest scores from 60 classes in three urban high schools. The results of standardized tests for evaluating students' attitudes toward their family indicate no significant curriculum effects for either unit. The cognitive test results are strong and unambiguous in all schools despite the differences in socioeconomic characteristics and school climate. Each unit did increase the students' knowledge of family life and their inquiry skills. (Author/DE)

ED107578

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

JUN 5

# EXPERIMENTAL EVALUATION OF FAMILY CURRICULUM MATERIALS FOR HIGH SCHOOL STUDENTS*

Shirley S. Angrist
School of Urban and Public Affairs

Anthony N. Penna
Department of History

Richard Mickelsen
School of Urban and Public Affairs

Carnegie-Mellon University

February 1975

SP 008 317

*This work was made possible by a grant from the Buhl Foundation. This paper will be presented at the American Educational Research Association annual meeting Washington, D.C. April 1, 1975.

00002

During the last decade major curriculum development projects in the social studies have assumed the importance of systematic formative evaluation. While varied in quality, such formative evaluations served to uncover the strengths and weaknesses in innovative curriculum materials and to assist developers with the task of revision.

Upon close scrutiny, a number of weaknesses emerge from the evaluation procedures employed by some of the major social studies development projects. Some projects engaged in no empirical evaluation of the curriculum materials. As a result, intuitive assessments of the successes and shortcomings of the project materials permeate their reports. Others employed limited empirical approaches, namely site visitations by developers, interviews and questionnaires (Far West Laboratory, 1969). Only a few of the thirty projects in the social sciences and history attempted summative evaluation through use of experimental research designs.

Edwin Fenton and his colleagues at Carnegie-Mellon University evaluated their experimental social science curriculum using the following procedures. They selected students on the basis of Otis IQ scores and teacher recommendations and then randomly divided them into control and experimental groups. After both groups took a full three years of social studies courses, the C-MU group administered a post test designed by the project's staff to measure student achievement of inquiry skills. "Post-experiment testing showed that students who followed the experimental program developed inquiry skills to a significantly greater degree than students in the control group" (Good, 1969).

The evaluation of the senior high school course, titled American Political Behavior by Mehlinger and Patrick (1971) represents another effort to

assess the effects of new materials on student learning. Tests for political knowledge, political science skills and political attitudes were developed by the project's staff. These measures were administered to students randomly divided into experimental and control groups in eight of the nine communities included in the study. The post test results revealed that students in experimental classes achieved higher scores on the political knowledge test than students in control classes in all eight communities. Mixed results appeared from the analysis of data on the political science skills test. In four communities, students in experimental classes scored significantly higher than control groups; modest to meager differences appeared in the remaining four communities. In the area of political attitudes, the treatment variable American Political Behavior had little or no effect on students (Patrick 1972).

The High School Geography Project (Kurfman, 1968) evaluated each of its one-month units with matched groups of students in experimental and control classes several times from 1965 through 1968. Verbal aptitudes of students in both groups and the teacher's knowledge of the discipline of geography were obtained by the Cooperative School and College Ability Test (SCAT) Form 3B and the GRE Advanced Geography test, respectively. A single unit test to measure student knowledge of geographic facts, generalization and concepts, and geography skills was administered to both experimental and control groups before and after the unit was taught. Results suggested that students in each successive year of field trials knew somewhat more about the content of each revised course than students in previous years.

Sociological Resources for Secondary Schools, a curriculum materials development project sponsored by the American Sociological Association (1968),

engaged in a systematic evaluation of eight of its projected twenty-seven episodes in spring 1967. We assume that the evaluation of these additional two-week episodes followed a similar research design. For evaluation purposes, episodes were paired. Students studying episode A took as a pre-test the 40 item post-test which would be administered to students studying episode B. Likewise, students studying episode B took as a pre-test the post-test which would be given to students taking episode A. Thus, an experimental group in one school served as a control group for an experimental group in another school. The SRSS evaluators collected the pre and post test results and matched the experimental and control groups according to variables such as verbal ability (obtained by administering The Psychological Corporation Verbal Abilities Test), socioeconomic status, and sex. The analysis of results for the five items which measured sociological perspeccive for each of the eight episodes suggested notable gains in three questions, but not much gain in two questions. On the remaining thirty-five episode-specific items, the gain in scores for experimental over control groups ranged from four to twelve percent (Far West Laboratory 1969).

These few curriculum projects stand out for their systematic experimental evaluations. Thus the conclusion reached by Mehlinger and Patrick (1971) whose own work is exemplary as systematic evaluation continues to have an authentic ring: "One feature of the 'new social studies' that has received less attention than it deserves is the place of evaluation in the special projects."

This paper represents a fresh attempt to systematically develop and evaluate innovative social studies materials. It describes the family life curriculum development project at Carnegie-Mellon University and presents the results of evaluation of two mini-courses using a modified Solomon four-group experimental design.

The Family Life Curriculum Project

In spite of the comprehensive curriculum development work completed by the various social studies projects in the 1960's, almost no systematic attempt has been made to write curriculum materials which focus on the sociological and historical development of the family. Until recently, secondary schools studied family life as a narrowly defined and relatively unimportant topic offered perfunctorily through home economics courses. This approach ignored the growing body of historical, psychological, sociological and anthropological research on the family, child development, child rearing, male and female roles, courting, mate selection and marriage. The apparent neglect of these topics in the high schools stemmed partly from a conviction that family life is too private a matter for the classroom. Furtner, prestigious social scientists for many years considered family life barely worthy of scholarly concern. Certainly, they doubted the ability of others to translate their research into instructional materials for secondary school students. Hence, little curriculum development has been done in this area.

The Social Studies Curriculum Center at Carnegie-Mellon University developed and evaluated two of its projected four mini-courses in the Family Life Curriculum.[1] Each course takes from about six weeks to teach and contains a student test, a teacher's guide, an audio/visual component, dittoed handouts for students, and tests. Beginning in fall 1973, two candidates for the Doctor of Arts degree in history, Stuart Lazarus and Mark Tierno, working with the project directors began to write curriculum materials for two of these courses: "American Youth: Past and Present" and "Family Life in Two Societies: Japan and the Kibbutz." In both courses, students employ

a conceptual structure referred to by developmental sociologists as the family life cycle and examine the changing norms, roles, role clusters, and socialization processes of family members in historical and contemporary contexts. Students learn these concepts along with inquiry skills to classify and analyze evidence about the topic. Vivid primary sources in the form of documents, letters, diaries, statistics, sketches, drawings and photographs are used as curriculum materials to link abstract concepts to information.

In addition, each course requires students to compare the experiences of family members in different historical periods. For example, in "American Youth: Past and Present," students investigate adolescent experiences in the eighteenth and nineteenth centuries as well as in the present. Searching for continuity and change, students are frequently asked to link the experiences of American youth in past time to their own. In the last chapter of this course, students learn about the adolescent experiences of their parents who most likely "grew up" in the late 1930's and the 1940's by examining materials contained in a series of data packets. Questionnaires are available for students to use as guides during discussions with their parents about their own adolescence. Students and teachers can create opportunities for parents to attend classes and participate in discussions in order to foster student learning. Each course aims to help students function as knowledgeable, thoughtful, and sensitive members of families headed by their own parents and to cope with role conflicts within the family.

"Family Life in Two Societies: Japan and the Kibbutz" has the added feature that students make cross-cultural as well as cross-generational comparisons. Both courses emphasize that the family is a universal social

institution which takes on variable forms according to historical and cultural contexts. The course materials promote natural rather than contrived connections to the daily lives of students.

The Experimental Design

The field trial of both courses was conducted in 1974 in four schools in the greater Pittsburgh region. Measures were utilized to assess knowledge about the family, ability to use inquiry skills, and attitudes toward the family through a modified Solomon four group design. The evaluation was designed to answer two separate but related questions:

1. To what extent do students experiencing either "American Youth: Past and Present" or "Family Life in Two Societies: Japan and the Kibbutz" learn about the family using particular inquiry skills?
2. To what extent do students experiencing either course acquire particular attitudes about the family in general and their own in particular?

The four schools in this study are located in three communities. School A is one of the largest public high schools in Pennsylvania and is located in a satellite city close to Pittsburgh. School B serves a high school age population in a working class industrial suburb of Pittsburgh. Schools C and D are located in an elite residential suburb of the city. The socio-economic characteristics of the students were obtained from a personal information questionnaire administered during the pre-test and post-tests (See Table 1). In Schools A and B, about 8 per cent and in Schools C and D, about 2 per cent were members of racial minorities, primarily black.

Fifty-five per cent of the students in Schools A and D and 45 per cent in School B and C were female. In religious affiliation between 47 per cent and 63 per cent were Catholic, but School B was most heavily Catholic, School A had about 34 per cent Protestant, Schools C and D had about 16 per cent each with "other religion" identified by five percent. In Schools A and B, fathers of students included in the study had attained less than a high school level education 33 per cent of the time. In Schools C and D, more than 65 per cent of the fathers had completed college. The students' self-report on academic performance indicates differences among the schools: students in Schools A and B received lower grades than those in Schools C and D.

In Schools A and B, students are permitted to enroll in social studies courses of their choice regardless of class rank or grade level. In School A, students select courses at the beginning of each semester or twice each year. In School B, the six week mini-course system operates and students choose new social studies courses six times each year. Of those enrolled in the two project courses in each of these schools, approximately 37 per cent were tenth graders, 38 per cent were eleventh graders, and 25 per cent were twelfth graders. Only in School B were students totally permitted to elect one of these courses from the mini-course offerings. Students in School A elected semester courses in sociology or youth culture but were not aware that the project materials would be included in these courses. In School C, across-grade registration was not permitted. In this school, the course titled "Family Life in Two Societies: Japan and the Kibbutz" was inserted in the tenth grade World Cultures curriculum. School D, the only junior high school included in the study, had begun to replace more

traditional course offerings with new ones. "American Youth: Past and Present" fit well into a seventh grade course entitled Youth Culture required at this school. That students in School B did elect to participate in the experiment and that students in School A elected a similar course with an assigned teacher may suggest a higher degree of motivation on the part of students in these schools to learn from the curriculum materials than from students in Schools C and D who did not either elect to participate in the experiment or request assignment to a comparison class. In both schools students were assigned to courses through typical scheduling procedures. This variation in designating students as experimental and comparison represents a design limitation of this study, since the experimenters could not randomly assign students to courses or to experimental groups. However, we found no evidence that experimental and control classes differed in any important way.[2]

Since all teachers volunteered to teach American Youth and Japan and Kibbutz, an additional limitation was placed on this study. Initially, their willingness to accept untried curriculum materials and to provide developers with feedback on the lessons implied enthusiasm, drive, and conscientiousness. But we discovered that some teachers may have been motivated more by a desire to receive printed student materials, daily teaching plans, an audio/visual component and ready-made tests than by an opportunity to participate in an experiment. Another variable separated the teachers in one school from those in the other three. Both developers were placed in School B as two of the nine teachers involved in the experiment in that school. They functioned at this site as teachers of their own materials, as in-service directors assigned the task of instructing six of

the school's social studies teachers in the proper use of these materials, and as supervisors of these teachers during the actual field-trial. Thus, School B received intensive in-service training. Schools A, C, and D received no similar preparation. A few classroom observations and site visitations before and during the field trial represented the only interventions in these schools. In all, 19 teachers in 56 classes participated in the experiment.

Since we could not randomly assign students to treatment and control groups, we refer to the research design as a "modified" Solomon four-group design. This design, according to Campbell and Stanley (1966), contains more empirical sources of internal and external validity than other experimental designs since it makes possible the use of analysis of variance to separate the effects of testing, the effects of curriculum and the interaction between testing and curriculum.

The experimental design appears in Table 2. The first group in this design receives pre-tests, then the curriculum and then is retested at the conclusion of the field trial. The second is a control group and takes the pre-tests, no curriculum materials, and post tests. By comparing the post test results of Groups 1 and 2, the effects of the curriculum materials are distinguished from the effects of maturation. Group 3 takes the post-test only. By comparing the post-test results for Groups 2 and 3, the effects of the pre-test are separated from the effects of maturation. Students in Group 4 receive the curriculum materials and the post-tests. Comparison of results for Group 4 and Group 1 separates the effects of the pre-tests from the effects of the curriculum.

Curriculum Impact on Attitudes

We assessed the effects of the two curriculum units on students' attitudes toward the family using three types of measures: a familism scale dealing with general conceptions about family, identification with one's own family and feelings about one's family.[3] The familism scale was adapted from The Family Scale developed by Sletto and Rundquist (1936). Examples of items are: home is the most pleasant place in the world; members of the family are too curious about one's personal affairs; so far as ideas are concerned, parents and children live in different worlds. The 21-item scale we used has responses ranging from 1 = strongly agree to 5 = strongly disagree, with the total score ranging from 21 to 105.

Identification with one's own family was tapped through two items developed by Simmons and Rosenberg (1973). This set had 2 questions which asked: If someone said something bad about your MOTHER, would you feel almost as if they had said something bad about you? This question was repeated referring to the respondent's FATHER. Responses were scored from 1 to 3 with a low score indicating high identification.

Another set of items tapped feelings about one's family: Do you feel that you are .... a very important, pretty important, not very important part of your family? How much do your parents care about the job you have when you get older? How much do you care about how good a (son) (daughter) you are? How good a (son) (daughter) are you? Responses to these four items were scored 1, 2 or 3 and added to form a score from 4-12, with a low score meaning higher identification.

Table 3 presents both pretest and post-test mean scores and standard deviations on familism for all schools taking each curriculum unit, distinguishing Group 1 which received the curriculum and Group 2 which did not. These results show only random differences between pre-test and post-test scores, with t-test values well below acceptable levels of statistical significance. While we do not present these results here, no changes from pretest to post-test occurred in identification or family feeling. Lack of gross differences with all schools combined and for both curriculum units and on all attitudinal measures suggested that further statistical analyses for the attitudinal measures were unwarranted.

From the lack of meaningful and statistically significant attitudinal differences among experimental and control groups, we draw several likely conclusions: (1) the measures may be insensitive to tap changes in perceptions of family life; (2) the measures could well be inappropriate and too remote since the curriculum materials did not directly raise issues related to familism; (3) six-week curriculum units may not affect students' perceptions of family life, rather longer exposure to such materials or different materials may be necessary. Given these tentative conclusions, we did not conduct any further analyses of the attitudinal measures.

Assessing the Curriculum Impact on Knowledge

While we used standard well-developed measures to tap the attitudinal areas, we developed our own tests for "cognitive" learning. To assess both the students' learning about the material covered and their facility to use inquiry skills, a test for each unit was constructed with items closely keyed to the unit's content and method. This strategy provided face validity of the items used.[4] In addition, to aid in developing valid tests, each of the

initial forms was tested with classes of high school students and college freshmen. These preliminary tests led to elimination of poorly worded items as well as to those which were answered correctly by very few or by nearly all respondents. The item analyses provide a third source of evidence of validity.[5] After item analysis, the test for American Youth contained 17 items and the test for Japan and Kibbutz had 22 items. Both tests consisted of multiple choice items with four response alternatives. Test scores computed as right minus 1/3 wrong (Cronbach, 1960).

We present the results on the cognitive tests first for all the classes in all the schools participating in the experimental evaluation. Next, we consider differences among schools and grade levels.

Results for All Schools

Test scores were analyzed using the method recommended by Campbell and Stanley (1966) for the four group Solomon design. The post-test scores were subjected to analysis of variance to sort out the effects of the curriculum, pre-testing, and their interaction effects. Tables 4 and 5 present the results for the Japan and Kibbutz unit, while Tables 6 and 7 show results for the American Youth unit.

The strong effects of the curriculum for both units are striking. As the F-ratios in Tables 4 and 6 show, effects of the pre-test and interaction effects were relatively small compared to curriculum effects. While Japan and Kibbutz had noticeably more impact than American Youth, it had some pretest effects. The mean scores in Table 5 highlight the curriculum effects: Groups 1 and 4 which received the curriculum have high scores and low standard deviations compared with Groups 2 and 3 which did not. The

results for American Youth differ somewhat. Seventh graders are affected by the curriculum and not by the pretest. The older students' scores reflect the two main effects both of the curriculum and the pretest.

While the curriculum effects are clear, it is important to note that pretest scores are rather high (see Tables 5 and 7). Thus it appears that even without the experimental curriculum units, some students are fairly competent in inquiry skills and have some knowledge about the family.

Results for Individual Schools: Japan and Kibbutz

When we consider each school separately, the strong curriculum effects again stand out for the Japan and Kibbutz unit. School C shows sizeable effects of the pretest; both Schools A and C exhibit interaction effects (see Table 8). School B, perhaps because the curriculum developers also taught there, shows the clearest uncontaminated effects of the curriculum, with no significant pre-test or interaction effects.

School A has generally high standard deviations and the lowest mean scores of the three schools, suggesting wide differences in student performance. This phenomenon confirms our personal observations that both students and teachers were less interested in the new material, and that daily discipline in this school was very lax. Compared with Group 3, the higher score for Group 2 indicates that the pretest may heighten the control group score slightly.

If we were to rank order the schools in terms of their mean scores, School C performs best overall, School B is next and School A last. This conclusion holds true both for control and experimental groups, but School C experimental groups performed notably better than the other two schools. This rank ordering probably reflects both school atmosphere and socioeconomic

differences. We already commented on the laxness of School A compared with the other two. School C draws from a high income area with highly educated families. The probable greater sophistication of these students on the esoteric subject matter of this unit is reflected in pretest and interaction effects. This is especially noteworthy because School C students are 10th graders while Schools A and B also include 11th and 12th graders. The pretest appears to depress test scores for both Schools B and C in Group 2 compared with Group 3; perhaps the pre-test control group shows its annoyance with repeated testing.

Results for Individual Schools: American Youth

For this unit, curriculum effects are also strong. This time both Schools B and D have the clearest results with the uncontaminated main effect of the curriculum (see Table 9). The contrast between School B's experimental and control groups is obvious in the two sets of mean scores. If we were to rank order the schools in terms of mean scores, School B would rank highest and School A lowest. Further, School D's generally good performance is impressive because these are all 7th grade students contrasted with grades 8-12 represented in the other two schools. But School D students are relatively high in socioeconomic status and that may partly explain their performance. This finding indicates that the American Youth unit is well suited even for the younger age group, particularly from higher socioeconomic backgrounds.

The strong main effect of the pre-test in School A is striking. Both for experimental and control groups (Groups 1 and 2), pre-testing depresses post-test scores, but this is truest for the control group. The pre-test

effect may reflect students' lack of interest and boredom with testing; many expressed annoyance to their teachers and to the investigators. But the general apathy of the teachers is also an important factor.

Taken together, the school results on the two units are significant. They underline the educational value of curriculum materials as short as six weeks in influencing students' knowledge about the historical, sociological and cross-cultural aspects of family life, as well as their ability to learn using inquiry skills.

## The Role of Teachers in the Experiment

For evaluation purposes, we were concerned with two matters: (1) the formative benefits derived from teachers' reactions to the materials and (2) the summative issue of whether teacher quality affected the impact of the curriculum units.

Teachers participated in the evaluation on a voluntary basis. We are satisfied that the study included a wide variety of teachers in terms of quality, experience, level of training, age, race and sex. It seemed equally clear that teachers varied greatly in their interest either in using the materials or in serving as "controls".

Teachers were helpful in raising questions and in showing weaknesses in the curriculum materials. We used their spoken and written comments to explain, clarify and correct some weaknesses whenever possible. However, not all the teachers wanted to spend time reporting their reactions -- we heard regularly only from a few teachers and prodded the others for comments. Mainly, teacher reactions were useful in revising the units <u>after</u> the evaluation.

Attempts to guage the effects of teacher quality had only limited success. We rated the 12 teachers who taught the units in three ways: (1) a priori, based on the investigators' impressions of their interest in and understanding of the materials,(2) student comments in their evaluation praising or deriding specific teachers and (3) a gain score reflecting the aggregated improvement of each class on the cognitive test between pre-test and post-test. Rank ordering teachers on these three criteria showed amazing incongruities -- only four of the twelve teachers ranked consistently on all three criteria. Two trends were evident: (1) Compared with teachers from the other schools more teachers from School B in which the curriculum developers taught showed high gain scores. Thus the daily meetings between the developers and teachers probably helped these teachers in reaching the students. (2) The two developers scored high both in gain scores and in student reactions, indicating that, compared with the other teachers, their superior knowledge of and enthusiasm for the material probably affected student learning.

Student Reactions to the Curriculum Materials

All students who participated in the post-test of the experimental groups were asked to complete an open-ended evaluation form. Their replies were used extensively to revise both units. Comments were content analyzed to indicate which aspects of the unit students most liked, least liked and to identify their suggestions for needed changes.

Students liked the Japan and Kibbutz curriculum mainly for their learning about other people, other cultures and life styles' and the possibility for historical and cross-cultural comparisons. Less often

mentioned but common were: satisfaction with the little work required and the few tests, group and individual participation, class discussion, good and easy-to-read materials, learning of concepts and overall positive feeling about the course. The least liked aspect was boredom with and repetition of concepts in class and in the text. Some students found the pace too slow, the material too easy; some disliked the cross-cultural focus, and others disliked the worksheets and data packets. Suggestions were made to raise the course level with more exciting, challenging material, projects and audiovisual aids as well as more class discussions.

The American Youth unit was most liked for its class discussions and student participation. In addition, students liked the variety, text and readings, newness of the material, openness of the class, audiovisual aids, stories of particular families and learning about the past. They least liked the large amount of writing, homework, worksheets, tests and data packets. Many also complained about the repetition of concepts in the class and textual materials. The students urged less written work and reading for this course with more discussion and more interesting materials. Some suggested more audiovisual aids and greater emphasis on contemporary material.

Since we hoped that learnings from the course might "spill over" into their own lives, we asked students several questions about these matters. As Table 10 shows, only some students thought the unit affected learning about their own family: thirty-five percent of those taking Japan and Kibbutz, and 23 per cent of those taking American Youth said so. Clearly, the spreading effects we hoped for did not occur for most students. This is particularly surprising for the American Youth course involving a familiar

context and culture. However, we cannot compare these reactions to other courses the students take and which are not labeled as "experimental."

Japan and Kibbutz students who felt this spillover did occur identified their learning as mainly about roles and norms, but also about feelings expressed at home and about historical changes and contrasts. American Youth respondents also frequently mentioned these kinds of learnings, but they emphasized insight into family situations and problems. Of those who answered "no", many felt they already knew these things, or that the course dealt with the past and with other families rather than with one's own family. Among American Youth students, many specifically commented that the course was irrelevant to one's own family. The fact that one quarter of the students were uncertain about their learning suggests difficulty in answering the question.

When we consider spillover in terms of discussing any part of the course with parents or schoolmates, the results are more encouraging. Discussion with schoolmates occurred considerably more than with parents, as might be expected for adolescents: 35 per cent of Japan and Kibbutz students and 42 per cent of American Youth students did so. More students from American Youth discussed the course than from Japan and Kibbutz; this may reflect the more immediate quality of a unit dealing with one's own age group (adolescents) in one's own country (See Table 11).

School differences are also noteworthy because they reinforce the cognitive test results. Students from the higher scoring schools and from higher socioeconomic levels (Schools C and D) discussed the material more both with parents and friends. The only exception occurred in School B where American Youth students who more often discussed the course with

parents than students from the other schools. This may reflect the strong family ties common to the ethnic groups represented in School B.

Summary and Conclusions

"American Youth: Past and Present" and "Family Life in Two Societies: Japan and the Kibbutz" are two of a projected four mini-courses in the Family Life Curriculum developed for secondary schools. The impact of these six-week units was assessed using an experimental design; at all times in the formal analyses the integrity of each of the two units was maintained, and thus what we report are really the results of two distinct experiments.

The results of the experiments for the two units are quite similar. In a sense, this is not a surprise; aside from the fact that two of the three schools in which each unit was tested are common between units, the units themselves bear a marked similarity. Both, of course, deal with family life, both employ the same conceptual structure defined by developmental sociologists as the family life cycle and examine the changing norms, roles, role clusters, and the socialization process of family members in several historical periods, including the present. Both units lasted about six weeks. Compared with other curriculum evaluations, it is striking to find such noticeable impact for courses of rather short duration.

There are, however, major differences between the units. Japan and the Kibbutz extensively employs cross-cultural comparisons in addition to the inter-generational comparisons of American Youth. American Youth deals with only a single culture, but with one experientially closer to the students. In the initial stages of the project, we planned to develop courses which would be inserted into typical ninth and tenth grade social studies and history courses. As we viewed it, American Youth would be taught in

ninth grade civics courses and Japan and Kibbutz would be taught in tenth grade World History or World Cultures courses. We selected high schools A, B, and C in order to measure the impact of the curriculum materials on the learning of students from different socioeconomic groups. Since Schools A and B had abandoned the conventional high school sequence of social studies and history courses, namely Civics, World Cultures, U.S. History and Problems of Democracy, and also eliminated age-segregation by abolishing grades ten, eleven, and twelve, we jettisoned our original plan for field trials in grades nine and ten. In addition, the inclusion of School D in the field trials provided us with an opportunity to measure the effects of American Youth on a junior high school population of seventh graders. We found that students in grades seven through twelve learn well using these courses. In order to stimulate upper grade high school students, however, the curriculum materials have been revised to increase their complexity by raising the reading level and including higher level cognitive activities.

Standardized tests for evaluating the students' attitudes toward family yielded no significant curriculum effects for either unit. The instruments were not designed specifically for the mini-courses, and thus some question of their appropriateness may be raised in retrospect; but the lack of curriculum-induced short-term attitude change is a common finding. Although most curriculum development projects in social studies have not attempted to measure change in student attitudes, the evaluation of American Political Behavior reported that no significant changes in political attitudes of students occurred as a result of taking the one-year course (Patrick, 1972). We are considering alternative approaches to measuring attitudinal changes.

Less formal evaluation techniques via an open-ended questionnaire given to those students who had received instruction in the experimental materials generally indicated that the cultural "breadth" of the Japan and Kibbutz unit was well received, while the "closeness" to adolescent experience postulated for the American Youth unit was viewed with mixed feelings. The purpose of the questionnaire, however, was that the results be used to improve the units. Indeed, the units have since been revised in light of student and teacher reactions, and the experimental results. Nonetheless, we take the frequency of discussion of the course materials outside the classroom and the demand for the advanced course materials as indicative of good reception.

The cognitive test results are strong and rather unambiguous in all the schools despite the differences in socioeconomic characteristics and school climate. The curriculum effects are highly significant and, in all but one instance, strongly dominate pre-test and interaction effects. Given that the two units were tested in schools having students of widely differing socioeconomic backgrounds and under many different teachers, we conclude that the six-week mini-courses did indeed increase the students' knowledge of family life and their inquiry skills.

Ideally, however, we would like both the curriculum materials and the cognitive tests to be more difficult and challenging, so that students start with the opportunity to achieve greater gains due to the curriculum. In addition, the tests should include more items in order to increase their reliability.

A combination of factors not included in the experimental design appear to have had some effect on the relative differences in results among the schools. The supervision afforded to the teachers in School B in using the new materials, and the general school environment in School A in terms of discipline are two noteworthy factors. While post-experiment attempts to identify the impact of these factors yielded no statistically significant rank order correlations, our belief that these affected the results, at least in terms of pre-test and interaction effects, remains.

"American Youth" and "Japan and the Kibbutz," as two mini-courses in the Family Life Curriculum, have been subjected to close scrutiny using formative and summative evaluation strategies. We conclude that the two approaches to curriculum development and evaluation are necessary and complementary. The results show that the curriculum units are effective.

FOOTNOTES

1. Two other courses are now being developed: "Courtship and Marriage: The American Experience" by James J. Paces and "Perspectives on the American Family" by Gerald M. Clarke.

2. We compared experimental and control classes on the following variables: year in school, grades, religion, father's education, race and sex. No statistically significant differences were found for any of these variables except sex. On the Japan and Kibbutz unit, School B had more girls in the experimental groups and for the American Youth unit, School A had more girls in the experimental groups than in the control groups. This may be due to the opportunities for students in Schools A and B to select courses and for girls to favor family related topics. Although girls are generally better academic achievers than boys especially in elementary school and thus could load the results for the experimental groups, this interpretation is not supported by other evidence: control and experimental classes are similar in grades and socioeconomic status, both of which reflect academic achievement.

3. All the instruments used in this study may be obtained by writing the authors.

4. We also analyzed the effects of subsections of each test separately by school and together for all schools. The subsections may be scored individually to distinguish content knowledge from inquiry skills. These results generally parallel those of the whole tests and are available upon request from the authors.

5. Item analyses were performed using pre-test data for Groups 1 and 2. The item pool was subsequently further reduced. Items retained for further analyses are characterized by an index of difficulty between 0.2 and 0.7, an index of discrimination above 0.25 (both based on an upper and lower 27 percentile comparison) and a point biserial correlation, adjusted for self-correlation, above 0.3 (Ebel, 1965; Thorndike, 1971).

We assessed reliability both in terms of equivalence and stability. One kind of evidence for equivalence comes from comparing pre-test scores for experimental and control groups, Groups 1 and 2. Tables 5 and 7 show that the means and standard deviations for the pretest experimental and control groups in the two curriculum units are very similar. We also determined "rational" equivalence using the Kuder-Richardson method for estimating reliability coefficients (Wert, 1954). For Japan and Kibbutz, the reliability coefficient was .76 and for American Youth .67. We assessed stability of the tests by comparison of mean scores and variances for the pre-test in Groups 1 and 2 with the post-test in control Group 3. For both curriculum units, some differences exist between the scores for these groups, but these differences are small and when examined on a school by school basis appear to be random.

## REFERENCES


American Sociological Association, 1967. National Trial of The Episodes: Family Roles, Social Mobility, Hypothesis Testing in The Social Sciences.

American Sociological Association, 1968. A Guide For Teachers Participating in National Trials of SRSS Episodes.

Campbell, Donald T. and Julian C. Stanley, 1963. Experimental and Quasi-Experimental Designs for Research. Chicago.

Cronbach, Lee J. 1960. Essentials of Psychological Testing. Harper and Row, New York.

Ebel, Robert L. 1965. Measuring Educational Achievement. Prentice-Hall, Englewood Cliffs, New Jersey.

Far West Laboratory for Educational Research and Development, 1969. "Integrated Information Unit: Secondary Social Studies Programs."

Good, John M., J. V. Farley and E. Fenton, 1969. "Developing Inquiry Skills with an Experimental Social Studies Curriculum." The Journal of Educational Research 63: 31-35.

Kurfman, Dana, 1968. "The Geography of Cultural Change" pp. 1-23 and "Growth of Cities" pp. 1-80: Two Evaluation Reports of the High School Geography Project. Association of American Geographers.

Mehlinger, Howard D. and John J. Patrick, 1971. "The Use of Formative Evaluation in an Experimental Curriculum Project." Social Education 35: 884-887, 892.

Patrick, John J. 1972. "The Impact of an Experimental Course, 'American Political Behavior,' on the Knowledge, Skills, and Attitudes of Secondary School Students." Social Education 36: 168-179.

Rundquist, E. A. and R. E. Sletto. 1936. Personality in the Depression: A Study in the Measurement of Attitudes. University of Minnesota Press, Minneapolis.

Simmons, Roberta G., F. Rosenberg and M. Rosenberg. 1973. "Disturbance in the Self-Image at Adolescence." American Sociological Review 38: 553-568.

Thorndike, Robert L. (Ed.), 1971. Educational Measurement. American Council on Education, Washington, D.C. (Second edition).

Wert, James E., C. O. Neidt and J. S. Ahmann. 1954. Statistical Methods in Educational and Psychological Research. Appleton-Century-Crofts, New York.

Table 1. Percent of Students in Four Schools With Selected Characteristics

| | Grade Level | | | | | |
|---|---|---|---|---|---|---|
| | 7 | 8 | 9 | 10 | 11 | 12 |
| School A | 0 | 1 | 0 | 37 | 40 | 23 |
| B | 0 | 0 | 0 | 38 | 36 | 26 |
| C | 0 | 0 | 1 | 97 | 1 | 1 |
| D | 99 | 1 | 0 | 0 | 0 | 0 |

| | Grades | | | |
|---|---|---|---|---|
| | Mostly A's | Mostly B's | Mostly C's | Mostly D's |
| School A | 15 | 44 | 38 | 2 |
| B | 9 | 41 | 45 | 5 |
| C | 14 | 60 | 23 | 3 |
| D | 23 | 61 | 14 | 1 |

| | Religion | | | |
|---|---|---|---|---|
| | Protestant | Catholic | Jewish | Other |
| School A | 35 | 54 | 2 | 9 |
| B | 27 | 63 | 1 | 9 |
| C | 29 | 49 | 7 | 16 |
| D | 27 | 47 | 9 | 16 |

| | Father's Education | | | |
|---|---|---|---|---|
| | Less than High School | High School | Some College | College |
| School A | 33 | 41 | 14 | 12 |
| B | 32 | 50 | 12 | 6 |
| C | 6 | 13 | 15 | 66 |
| D | 1 | 17 | 12 | 70 |

| | Percent Minorities | Percent Female |
|---|---|---|
| School A | 7 | 55 |
| B | 8 | 44 |
| C | 2 | 41 |
| D | 2 | 54 |

Table 2. The Solomon Four Group Experimental Design

| | | | |
|---|---|---|---|
| Group 1 | Pretest | Curriculum | Post-test |
| Group 2 | Pretest | -- | Post-test |
| Group 3 | -- | -- | Post-test |
| Group 4 | -- | Curriculum | Post-test |

Table 3. Comparison of Pretest and Post-test Scores on Familism*

| | | Mean | Standard Deviation | N |
|---|---|---|---|---|
| Japan and Kibbutz: | | | | |
| Group 1 | Pretest | 64.4 | 11.6 | 177 |
| | Post-test | 62.4 | 10.9 | 95 |
| Group 2 | Pretest | 62.4 | 10.9 | 187 |
| | Post-test | 63.9 | 11.5 | 132 |
| American Youth: | | | | |
| Group 1 | Pretest | 61.8 | 10.5 | 149 |
| | Post-test | 62.4 | 11.0 | 140 |
| Group 2 | Pretest | 60.7 | 10.6 | 156 |
| | Post-test | 60.3 | 10.8 | 101 |

*None of the t-tests between pretest and post-test mean scores indicated significant differences.

Table 4. Analysis of Variance for All Schools on the Japan and Kibbutz Curriculum

| Source of Variance | DF | Sum of Squares | F-ratio |
|---|---|---|---|
| Curriculum | 1 | 1763.73 | 188.35** |
| Pretest | 1 | 39.11 | 4.18* |
| Interaction | 1 | 30.72 | 3.28 |
| Between | 3 | 1858.77 | |
| Within | 419 | 3923.60 | |
| Total | 422 | 5782.38 | |

*$p < .05$

**$p < .01$

Table 5. Pre-test and Post test Mean Scores for Four Experimental Groups in All Schools on the Japan and Kibbutz Curriculum

| Group | Pre-test | | | Post-test | | |
|---|---|---|---|---|---|---|
| | Mean | S. D. | N | Mean | S. D. | N |
| 1 | 16.8 | 3.3 | 163 | 19.1 | 2.6 | 104 |
| 2 | 16.1 | 3.1 | 189 | 14.5 | 3.3 | 136 |
| 3 | | | | 15.6 | 3.4 | 96 |
| 4 | | | | 19.1 | 2.6 | 87 |

Table 6. Analyses of Variance for 7th grade and 8-12th grades on the American Youth Curriculum

| 7th grade only: Source of variance | Degrees of Freedom | Sum of Squares | F-ratio |
|---|---|---|---|
| Curriculum | 1 | 212.75 | 50.55* |
| Pretest | 1 | 7.96 | 1.89 |
| Interaction | 1 | 12.15 | 2.89 |
| Between | 3 | 230.66 | |
| Within | 264 | 1111.00 | |
| Total | 267 | 1341.66 | |

| 8-12th grades: Source of variance | Degrees of Freedom | Sum of Squares | F-ratio |
|---|---|---|---|
| Curriculum | 1 | 244.27 | 40.50* |
| Pretest | 1 | 76.20 | 12.63* |
| Interaction | 1 | 9.26 | 1.54 |
| Between | 3 | 294.28 | |
| Within | 274 | 1652.62 | |
| Total | 277 | 1946.90 | |

*$p < .01$

Table 7. Pre-test and Post-test Mean Scores for Four Experimental Groups in All Schools on the American Youth Curriculum

| Group | Pre-test | | | Post-test | | |
|---|---|---|---|---|---|---|
| | Mean | S. D. | N | Mean | S. D. | N |
| 7th grade only: | | | | | | |
| 1 | 12.2 | 2.2 | 54 | 13.7 | 2.1 | 55 |
| 2 | 12.1 | 2.2 | 52 | 11.4 | 2.2 | 52 |
| 3 | | | | 12.2 | 2.1 | 83 |
| 4 | | | | 13.6 | 1.7 | 78 |
| 8-12th grades: | | | | | | |
| 1 | 11.5 | 2.1 | 104 | 12.7 | 2.6 | 90 |
| 2 | 11.7 | 2.2 | 103 | 10.5 | 2.4 | 62 |
| 3 | | | | 11.9 | 2.2 | 71 |
| 4 | | | | 13.4 | 2.4 | 55 |

Table 8. Summary of Analyses of Variance for Three Schools on Japan and Kibbutz Curriculum

| Group | School A (N=174) | School B (N=105) | School C (N=144) |
|---|---|---|---|
| 1. Mean | 16.7 | 19.1 | 20.4 |
| Standard deviation | (2.7) | (2.3) | (1.6) |
| 2. Mean | 14.1 | 14.8 | 14.8 |
| Standard deviation | (3.4) | (3.3) | (3.2) |
| 3. Mean | 13.7 | 16.4 | 18.5 |
| Standard deviation | (3.0) | (2.4) | (2.4) |
| 4. Mean | 18.5 | 19.0 | 20.6 |
| Standard deviation | (2.7) | (2.4) | (1.7) |
| F-ratio | | | |
| Curriculum | 59.71** | 41.54** | 122.68** |
| Pretest | 1.03 | 1.74 | 22.06** |
| Interaction | 5.06* | 2.77 | 15.56** |

*$p < .05$

**$p < .01$

Table 9. Summary of Analyses of Variance for Three Schools on American Youth Curriculum

| Group | School A (N=181) | School B (N=97) | School D (N=268) |
|---|---|---|---|
| 1. Mean | 11.6 | 14.4 | 13.7 |
| Standard deviation | (2.2) | (2.2) | (2.1) |
| 2. Mean | 10.2 | 11.4 | 11.4 |
| Standard deviation | (2.3) | (2.5) | (2.2) |
| 3. Mean | 12.1 | 11.6 | 12.2 |
| Standard deviation | (2.3) | (1.9) | (2.1) |
| 4. Mean | 12.4 | 15.4 | 13.6 |
| Standard deviation | (2.4) | (1.1) | (1.7) |
| F-ratio | | | |
| Curriculum | 6.87* | 59.49* | 50.55* |
| Pretest | 16.72* | 1.99 | 1.89 |
| Interaction | 2.62 | 0.82 | 2.89 |

*$p < .01$

Table 10. Do You Think You Learned Anything About Your Own Family?

| | Japan and Kibbutz (N=165) | American Youth (N=244) |
|---|---|---|
| Yes | 35 | 23 |
| No | 37 | 52 |
| Don't Know | 28 | 25 |
| Total | 100% | 100% |

Table 11. Discussion With Parents and Schoolmates

Japan and Kibbutz:
(N=165)

| School | Percent Discussed With Parents | Percent Discussed With Schoolmates |
|---|---|---|
| A | 13 | 24 |
| B | 18 | 39 |
| C | 22 | 42 |
| Total | 18 | 35 |

American Youth:
(N=244)

| School | Percent Discussed With Parents | Percent Discussed With Schoolmates |
|---|---|---|
| A | 27 | 27 |
| B | 35 | 42 |
| D | 30 | 52 |
| Total | 28 | 42 |